



RESEARCH DEPARTMENT



REPORT

**A digital split-band compandor
for 64 kbit/s coding of speech
with a 7 kHz bandwidth**

W.I. Manson, B.Sc.(Eng.)
D.W. Stebbings, B.A.(Cantab.)

BBC RD 1977/41
UDC 534.86:
621.376.56:
621.395.665.1

A DIGITAL SPLIT-BAND COMPANDOR FOR 64 KBIT/S
CODING OF SPEECH WITH A 7 KHZ
BANDWIDTH

W.I. Manson, B.Sc.(Eng.)
D.W. Stebbings, B.A.(Cantab.)

Summary

This Report describes a three-channel split-band digital compandor arrangement for coding a 7 kHz speech signal for transmission at 64 kbit/s. It was not practicable at the time to construct equipment for real-time evaluation, but the performance of a simulation has been assessed. The results obtained are regarded as promising, but the performance was partially obscured by imperfections unavoidably introduced in the simulation. It is recommended that a real-time processor be built to allow the true potential of the system to be assessed.

Issued under the authority of



Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

Head of Research Department

November 1977

(EL-129)

A DIGITAL SPLIT-BAND COMPANDOR FOR 64 KBIT/S CODING OF SPEECH WITH A 7 KHZ BANDWIDTH

Section	Title	Page
	Summary	Title Page
1.	Introduction	1
2.	Digital companding	1
	2.1. General	1
	2.2. Split-band digital companding	1
	2.3. Experimental 64 kbit/s apparatus	2
3.	Other coding arrangements assessed	3
4.	Subjective investigation	3
	4.1. Test arrangements	3
	4.2. Results of subjective investigation	3
5.	Conclusions	5
6.	References	5

A DIGITAL SPLIT-BAND COMPANDOR FOR 64 KBIT/S CODING OF SPEECH WITH A 7 KHZ BANDWIDTH

W.I. Manson, B.Sc.(Eng.)
D.W. Stebbings, B.A.(Cantab.)

1. Introduction

The background to the present interest in the coding of speech for transmission in 64 kbit/s digital channels has been described fully in earlier work.^{1,2} Briefly, the situation is that, while analogue sound circuits are likely to become progressively less readily available, the development of a telephone network based on a 64 kbit/s data channel — planned internationally by Post Office organisations — should ultimately give broadcasters direct access to a very widespread 64 kbit/s network which could be used for commentaries and speech contributions by correspondents, both nationally and internationally.

However, the speech coding parameters for the digital public telephone service restrict the channel bandwidth to some 3.4 kHz — about half the bandwidth which is widely regarded as desirable for broadcast speech. There is therefore considerable incentive to devise a bit-saving technique to enable good-quality, 7 kHz bandwidth, speech signals to be coded for transmission in a 64 kbit/s channel. A number of bit-saving techniques have already been investigated.² Some were considered promising, giving results which approached the required standards; however, side effects were produced which were sometimes obvious, and it was concluded that a better solution should be sought.

The present work describes a further method of bit saving to enable a 7 kHz speech signal to be transmitted at 64 kbit/s. It also gives the results of a subjective evaluation both of this new system and also of some of the systems described previously.

2. Digital companding

2.1. General

The most widely applied technique for bit-rate reduction of digital sound signals is that of digital companding. This is, essentially, a non-linear quantising process in which the quantising steps are small for low-level signals, where small steps are necessary to avoid the so-called granular or crumbling distortion, and progressively larger for the higher-level signals where the consequent increase in quantising noise is more likely to be concealed, or masked, by the presence of the signal itself. The near-instantaneous digital companding technique,³ for example, enables the bit rate for high-quality sound signals to be reduced by about 25% below that required for linear coding, without significant loss of quality of the reproduced programme.

An even greater bit-rate reduction is required for the present purpose. However, in a non-linear coding system, the increased noise will not be adequately masked

by the programme if the quantisation step-size for large signal levels is too great. The resulting programme-modulated ('hush-hush') noise can be particularly obtrusive in a full-band compandor system when noise components which lie in a frequency region where the ear is sensitive are modulated by signal components which lie in a frequency region where the ear is less sensitive and which therefore provide less-effective masking. The modified companding technique to be described operates by giving improved masking and, further, by reducing the effective level of quantising noise.

2.2. Split-band digital companding

The degradation of quality due to programme-modulated noise introduced in companding can be reduced by dividing the audio-frequency range into a number of bands which are then companded separately — a technique which has been applied successfully to analogue companders.⁴

In the first place, split-band companding restricts the components of the hush-hush noise from each compandor to the frequency band containing the signal components which gave rise to them; as a result of this close proximity of the noise and signal spectra, masking is considerably enhanced.

Further, band splitting results in a reduction in the effective level of programme-modulated noise in the following manner. In companding the level of programme-modulated noise is related to the level of the applied programme signal; any arrangement which reduces the level of signal also reduces the level of programme-modulated noise produced. Where the audio frequency spectrum is split into bands before companding, the signal level applied to each compandor — and, therefore, the amount of programme-modulated noise produced in any band — is less than in a corresponding full-band system. This effect is particularly important since the ear's sensitivity is high at medium to high audio frequencies, yet the energy in this part of the spectrum is usually relatively small. Thus programme material which, on a wideband compandor system, might well cause audible impairment due to programme-modulated noise, might rarely activate the top band of a split-band companding system.

In considering a practical system for a split-band digital compandor it became clear that the larger the number of bands into which the audio spectrum is divided the greater is the potential for bit-saving. However, the complexity of the apparatus required increases correspondingly — separate, or time-division-multiplexed, coders and decoders are required to equip each channel, together with the necessary band-splitting filters and the appropriate digital logic units. A practical com-

promise was chosen for the experimental compandor, dividing the programme spectrum into three bands — up to 1.75 kHz, 1.75 kHz to 3.5 kHz and 3.5 kHz to 7 kHz.

The choice of frequency limits for each band is particularly significant in a bit-saving system of the kind being considered. The selection of limits extending from f to $2f$ for the two upper bands enables them to be coded with a sampling rate of one-half that which might, at first sight, appear to be necessary. This may be achieved by either of two alternative techniques. In one, the signals in each band are translated to baseband before coding, and restored to the original form after decoding. Thus the 1.75 kHz to 3.5 kHz band is translated to 0 to 1.75 kHz and the 3.5 kHz to 7 kHz band to 0 to 3.5 kHz, and sampling can be carried out at about 3.5 kHz and 7 kHz respectively. In the second technique advantage is taken of the fortunate disposition of the alias components which are produced when the f to $2f$ bands are each sampled virtually at its upper frequency. Fig. 1 shows the spectrum of the sampled signal; there is no overlapping of the sidebands generated in the sampling process, and the wanted signal can be recovered by band-pass filtering of the decoded signal reconstituted using narrow pulses, rather than a box car waveform, so as to maintain the response up to the sampling frequency. None of the sidebands generated by such sampling falls within the octave bandwidth of the original signal, so no alias components are included in the reconstituted signal. Other workers have independently considered similar artifices to reduce the sampling rates necessary for band-split digital systems.⁵ For the present investigation frequency-shifting equipment was already available, and the technique of translating the two upper bands to baseband before coding was therefore adopted.

The transmitted data rate for the investigation has been assumed to be 64 kbit/s. This can be allocated to the three bands in various ways — improving the performance in some, while degrading it in others. The 0 to 1.75 kHz and 1.75 to 3.5 kHz bands are obviously important for natural and intelligible speech quality and 6-bit words were allocated to each, requiring a combined data rate of 42 kbit/s. The band from 3.5 to 7 kHz, which, in speech,

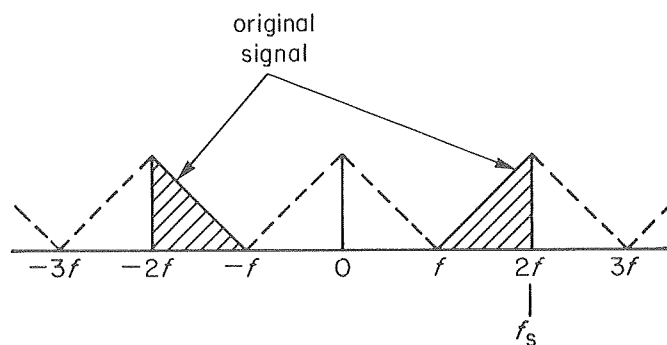


Fig. 1 - Spectrum of f to $2f$ signal, sampled at $2f$

contains relatively little energy, was allocated 3 bits/word. With a sampling rate of 7 kHz this band also required a data rate of 21 kbit/s, making a total data rate of 63 kbit/s for the three bands and leaving 1 kbit/s for ranging information for the three compandors.

Table 1 lists the parameters for the arrangement selected for practical study.

2.3. Experimental 64 kbit/s apparatus

In the absence of sufficient equipment for complete signal processing, the performance of the system described above was assessed by simulation, using a multiple-recording technique. The channels were processed separately and then recombined to restore the original 7 kHz signal. Each channel was derived by appropriate filtering, frequency shifted (in the case of the upper two bands) as described above, coded at the appropriate sampling rate, subjected to the multi-range near-instantaneous companding law indicated, and then recovered by a complementary process.

Although processing was not carried out in real time, a reasonable evaluation of the system was none-the-less possible. However, the performance of the simulation was undoubtedly pessimistic, partly as a result of degradation due to the multiple dubbing required to build up the final test tape, and partly because the available filters used in the experiment were not ideal and allowed some unwanted tones and intermodulation products to be present in the

TABLE 1

Parameters of experimental split-band near-instantaneous companding system

Frequency Band	Sampling Rate	Transmitted Bits/Sample	No. of Ranges in Companding Law	Bits/Scale Factor	Scale Factor Period mS	No. of Samples Delay	Scale Factor Bit-rate bit/s
0 — 1.75 kHz	3.5 kHz	6	8	3	10	35	300
1.75 — 3.5 kHz	3.5 kHz	6	8	3	10	35	300
3.5 — 7.0 kHz	7.0 kHz	3	11	4	10	70	400

simulation. Some indication of the degree of pessimism was obtained by including in the subjective investigation items which had undergone the filtering, multiple-dubbing and frequency-shifting processes described above, but not the digital coding, companding and decoding; ideally this analogue processing should cause no quality degradation.

3. Other coding arrangements assessed

Two other bit-saving systems, which were considered earlier, were included in the present assessment for comparison. Full details of these systems have been given previously² but, for convenience, they are briefly outlined below.

(i) "Pitch-halving" technique

Briefly, this technique relies on the fact that, in a speech waveform, adjacent short segments of, say, 16 ms duration are often very similar to each other. Transmission of alternate segments only — expanded in time so that both the bit rate and the pitch are halved — may then suffice, provided that the transmitted segments, after being restored to the original pitch, are each presented twice by the receiving terminal to provide continuity of signal. Pre- and de-emphasis, together with comb-filtering, were introduced in the original work to mitigate some of the more obvious imperfections of the system, and were included also in the present simulation.

(ii) "Variable-sampling-rate" technique

In this arrangement the sampling-rate is varied automatically according to the needs of the programme signal at the time, the number of bits per word being changed in sympathy to keep the transmission bit rate constant. Thus the number of bits per word is in general kept as high as is possible while the level of quantising noise is kept down.

A simple conventional coding arrangement having a 16 kHz sampling rate, 4-bits/word, and 10-range near-instantaneous companding, was also included for comparison. This was not expected to compete with the systems described earlier, but is of interest as it provides a 7 kHz channel encoded at 64 kbit/s using an adaptation of a technique which is well-established for 15 kHz high-quality sound coding.

4. Subjective investigation

4.1. Test arrangements

The main purpose of the investigation was to assess the merits of the experimental coding system for speech transmission; however, because of encouraging preliminary results, a number of musical items were also included. Tests included excerpts of two male voices, one female

voice, solo piano, solo trumpet, light music and orchestra.

The selected excerpts were processed as follows:

Coding System	Method
1. Pitch-halving technique	Simulated
2. Variable-sampling-rate technique	Simulated
3. Split-band digital companding technique	Simulated
4. As 3 but excluding digital processing	*
5. Full-band 10-range near-instantaneous companding. 4 bits/sample 16 kHz sampling rate	Real time

Recordings of the various test items, unprocessed apart from a 7 kHz band restriction, were also included for reference.

The test tape was arranged so that presentations of each programme item were grouped together, with unprocessed programme presented first, for reference, and the other processing systems presented thereafter in random order.

Thirteen listeners, mostly experienced in assessing sound programme quality, attended the test sessions in groups of up to six. They were asked to grade the presentations of the processing systems (1 to 5) according to the CCIR 5-point impairment scale given below:

1. Very annoying
2. Annoying
3. Slightly annoying
4. Perceptible but not annoying
5. Imperceptible

4.2. Results of subjective investigation

Table II gives the mean gradings for the various presentations, together with the corresponding values of standard error.

If we consider first the tests carried out with speech, it is apparent that System 5, using 16 kHz sampling and 4 bits/word with full-band companding, is, as might be expected, unacceptable; the mean scores obtained were all below Grade 2. System 3, the split-band digital companding arrangement, obtained mean scores between Grade 3 and Grade 4; however, it is of particular interest that System 4, comprising simply the analogue processes of System 3, itself degraded the programme to about Grade 4. The way in which the impairments due to the analogue and digital effects combined in System 3 is not known, but it is evident

* Comprising the analogue multiple-dubbing, filtering and frequency-drifting processes used in the simulation of System 3.

TABLE II

*Results of subjective tests: mean gradings
(standard errors given in italics)*

Speech

Processing system	Programme item					
	Speech Male (1)	Speech Male (2)	Speech Female		Speech (Overall)	
1. Pitch-halving	2.85 <i>0.17</i>	2.65 <i>0.32</i>	2.15 <i>0.21</i>		2.55 <i>0.14</i>	
2. Variable-sampling rate	2.81 <i>0.19</i>	2.12 <i>0.20</i>	2.42 <i>0.19</i>		2.45 <i>0.12</i>	
3. Split-band digital companding	3.42 <i>0.18</i>	3.73 <i>0.24</i>	3.19 <i>0.16</i>		3.45 <i>0.11</i>	
4. As 3, but excluding digital processing	4.19 <i>0.13</i>	3.96 <i>0.19</i>	4.0 <i>0.21</i>		4.05 <i>0.10</i>	
5. Full-band N.I. companding, 16 kHz sampling rate, 4 bits/sample	1.62 <i>0.17</i>	1.50 <i>0.19</i>	1.04 <i>0.04</i>		1.39 <i>0.09</i>	

Music

Processing system	Programme item					
	Solo Piano	Solo Trumpet	Light Music with Vocal	Orchestra		Music (Overall)
1. Pitch-halving	3.77 <i>0.23</i>	1.12 <i>0.08</i>	2.0 <i>0.25</i>	2.38 <i>0.17</i>		2.32 <i>0.16</i>
2. Variable-sampling rate	3.27 <i>0.28</i>	3.54 <i>0.31</i>	4.19 <i>0.20</i>	1.19 <i>0.11</i>		3.05 <i>0.19</i>
3. Split-band digital companding	4.0 <i>0.17</i>	2.5 <i>0.14</i>	3.19 <i>0.18</i>	3.65 <i>0.19</i>		3.34 <i>0.12</i>
4. As 3, but excluding digital processing	3.77 <i>0.17</i>	4.31 <i>0.18</i>	4.15 <i>0.15</i>	4.27 <i>0.20</i>		4.12 <i>0.09</i>
5. Full-band N.I. companding, 16 kHz sampling rate, 4 bits/sample	1.08 <i>0.08</i>	1.0 <i>0</i>	3.58 <i>0.26</i>	1.46 <i>0.13</i>		1.78 <i>0.17</i>

that the instrumental imperfections in the analogue equipment used in simulating System 3 must have degraded the overall performance to some extent, and that an actual split-band digital compandor would give a better performance than that achieved by the simulation.

Systems 1 and 2 gave very similar results, all lying between Grade 2 and Grade 3. They thus give consistently worse performance than System 3.

The results for the musical test items were more scattered, and it appears that some types of music are

more tolerant, and others less tolerant, of particular processing systems. However, the conclusions to be drawn are in general the same as for speech tests. System 5 is unacceptable. System 3 generally has a moderate grading although the results are again pessimistic to some degree because of the known limitations of the simulation technique used. System 1 generally obtained a low to moderate grading while System 2 obtained a fairly good grading for three items, but a low grading on the fourth item on which the adaptive changes in system bandwidth were particularly obtrusive.

5. Conclusions

There is an attraction in using the future national and international digital telephone networks to provide improved quality for broadcast telephone speech contributed by reporters and correspondents.

This Report has described a split-band digital compandor system devised principally to code 7 kHz speech signals into 64 kbit/s, the bit rate corresponding to a digital telephony channel. The performance of a simulation of the system has been assessed subjectively and the results are considered very promising, particularly when account is taken of impairment introduced as a result of known shortcomings in the analogue processes used in the simulation. Further effort is necessary to construct an actual split-band compandor system and to assess its full potential.

Though considered in the first instance for 7 kHz speech applications, split-band digital companding may also serve as an effective bit-saving technique for full-bandwidth high-quality sound signals. Any equipment constructed for further work on 7 kHz signals should be arranged so that it can also be extended for 15 kHz applications.

6. References

1. CROLL, M.G. 1975. The possibility of sending programme speech contributions digitally over the public telephone network. BBC Research Department Report No. 1975/6.
2. STEBBINGS, D.W. 1976. An experimental comparison of four methods for 64 kbit/s coding of speech with a 7 kHz bandwidth. BBC Research Department Report No. 1976/11.
3. OSBORNE, D.W. and CROLL, M.G. 1973. Digital sound signals: bit-rate reduction using an experimental digital compandor. BBC Research Department Report No. 1973/41.
4. DOLBY, R.M. 1967. An audio noise reduction system. J. Audio Eng. Soc. **15**, pp 383-388.
5. CROCHIERE, R.E., WEBBER, S.A. and FLANAGAN, J.L. 1976. Digital coding of speech in sub-bands. Bell System Technical Journal, **55**, No. 8, pp 1087-1109.

